

NPS ARCHIVE
1966
EARNEST, C.

ESTIMATING RELIABILITY AFTER CORRECTIVE
ACTION: A BAYESIAN VIEWPOINT

CHARLES M. EARNEST

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CA 93943-5101

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIF. 93940

ESTIMATING RELIABILITY AFTER CORRECTIVE ACTION:
A BAYESIAN VIEWPOINT

by

Charles M. Earnest
Lieutenant Commander, United States Navy
B. S., Auburn University, 1955

Submitted in partial fulfillment
for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

UNITED STATES NAVAL POSTGRADUATE SCHOOL
May 1966

NPS Archive
1966
Earnest, C.

~~Theo/s~~
~~E1/8~~
~~c1~~

RECEIVED
11-10-66
11-10-66

ABSTRACT

A complex system is considered in its latter stages of development. N mission trials have been observed, each resulting in a success or a failure. Each failure occurs in one of k failure modes. For each failure mode that is observed action is taken to attempt to correct that type of failure. The probabilities of correcting the various failure modes are known. After corrective action is completed attempts to estimate the current reliability, without further sampling, are made. A brief historical summary of this problem to date is given.

Justification for assuming a prior distribution on the failure modes is discussed and the posterior distribution of the parameters is developed. An intuitive measure of the current reliability is stated and certain properties of this random variable are developed.

TABLE OF CONTENTS

Section	Page
1. Introduction	5
2. Bayesian Viewpoint	10
3. Prior Distribution of Component Failures	12
4. Posterior Distribution of the Parameters	19
5. Reliability	21
Bibliography	26
Appendix I - Selection of Prior Parameters	27

1. INTRODUCTION

In the latter stages of development and testing of a complex system it is reasonable to assume that a certain number of mission trials have been effected, and of these some number resulted in the failure of the mission. An intuitive measure of the reliability of the system at this point would be the ratio of failures to the total number of tests conducted. After analysis of the failure data however, it is conceivable that the reasons for each failure of the mission could be detected and some corrective action taken against this type of failure. After taking corrective action the natural step would be to continue testing and make statements about the reliability of the revised system on the basis of new data.

If the cost of additional testing is high, if additional test units are not available, or if time is a prohibitive factor, further testing may not be feasible. At this point Corcoran, Weingarten and Zehna [1] pose the following problem: "Assuming that we have confidence in our knowledge of the effectiveness of the contemplated corrective action which may be taken on observed failure modes, how should we use the results of the first N tests to draw inferences about the current reliability?" The problem is specifically structured in this manner. Let N tests be conducted. Each of the tests results either in a successful performance of the mission or in a failure. Each failure results from a failure of one or k possible modes. The probability of a successful mission is q_0 , the probability of a mission failing by the i^{th} failure mode is q_i where i ranges from one to k . The q 's themselves are unknown; however, since each trial must result in a success of some

type of failure it may be noted that

$$\sum_{i=0}^k q_i = 1.$$

Similarly if we let N_i be the number of events of the i^{th} type observed where N_0 is the number of successful missions then

$$\sum_{i=0}^k N_i = N$$

It is also assumed that there is a known conditional probability a_i of correcting the i^{th} failure mode given that it occurs.

Based on the above formulation Corcoran, Weingarten, and Zehna [1] define the following random variable as a measure of current reliability:

$$p^* = q_0 + \sum_{i=1}^k y_i q_i \quad \text{where } y_i = \begin{cases} 0 & \text{if } N_i = 0 \\ a_i & \text{if } N_i > 0 \end{cases}$$

This is stated to be an intuitive measure of the reliability since it adds a weighted amount of the failure probability of each observed failure mode to the initial reliability. The expected value of p^* may be computed and is referred to as the "mean reliability." This quantity is shown to be:

$$E[p^*] = q_0 + \sum_{i=1}^k a_i q_i [1 - (1 - q_i)^N]$$

Since p^* is a random variable it is customarily not estimated. However, the variance of p^* tends to zero as N tends to infinity; therefore any estimate of $E[p^*]$ can be said to asymptotically estimate p^* . The above authors^[1] postulate seven estimators of $E[p^*]$ and discuss their relative

merits from a standpoint of bias. The last two estimators, quoted here for convenience, are conservative (underestimate $E[p^*]$) and are asymptotically unbiased. The last estimator, p_7 , is the more conservative of the two but is also consistent for $E[p^*]$.

$$p_6 = \frac{N_0}{N} + \sum_{i=1}^k z_i \frac{N_i}{N} \quad \text{where } z_i = \begin{cases} a_i & \text{if } N_i > 1 \\ 0 & \text{otherwise} \end{cases}$$

$$p_7 = \frac{N_0}{N} + \sum_{i=1}^k y_i \frac{(N_i - 1)}{N} \quad \text{where } y_i = \begin{cases} a_i & \text{if } N_i > 0 \\ 0 & \text{if } N_i = 0 \end{cases}$$

The bias for estimator p_6 is shown not to exceed .01 for an N of 25 and a q_0 as large as one third.

Larson [3] considers the conditional distribution of the reliability (conditioned on the outcome of the test) and demonstrates its probability mass function. Unfortunately while the mass function is known, the actual values the random variables take on are unknown since they are functions of the q_i 's. A functional lower bound on the true reliability is shown to be:

$$P[Z \geq f|A] \geq P_m[g \geq f|A] \quad P[Z \geq g|A]$$

where the conditioning event A is the event that $N_1 > 0, \dots, N_r > 0$, f is a function of the observable random variables and g is some function of the unknown parameters. $P_m[g \geq f|A]$ is a conditional multinomial probability statement and $P[Z \geq g|A]$ is a statement from the distribution of true reliability conditioned on sample results. The expression $P[Z \geq g|A]$ can be evaluated for a certain class of g functions¹; thus a

¹Larson, H. J., Conditional Distribution of True Reliability after Corrective Action (US Naval Postgraduate School. Technical Report/Research Paper No. 61, 1966) p. 10.

lower bound on the reliability can be obtained if the appropriate conditional multinomial statement can be derived.

This approach has been considered by the present investigator but abandoned due to the compounding complexity in the probability statements. The class of functions of the parameters (q_i 's) is restrictive since the $P[Z \geq g|A]$ must be capable of evaluation. No suitable functions of the observable random variables have been found that lead to direct evaluation of the multinomial statement. The conditional distribution of the sample can be obtained but it becomes exceedingly complex as the number of failure modes increases. As an example if we consider only three modes of failure (N_1, N_2 and N_3) and the conditioning event $A = \{N_1 > 0, N_2 > 0, N_3 = 0\}$ then the conditional probability is as follows:

$$P[N_0 = a, N_1 = b, N_2 = N - (a+b) | A]$$

$$= \frac{N! q_0^a q_1^b q_2^{N-(a+b)}}{a! b! [N-(a+b)]! [(q_0 + q_1 + q_2)^N - q_0^N - (q_0 + q_1)^N - (q_0 + q_2)^N]}$$

$$\text{for } \begin{cases} 0 < a < N-2 \\ 0 < b < N-a \end{cases}$$

$$= 0 \quad \text{otherwise}$$

It would seem that any direct attempt to derive the distribution of some function of the N_i 's is hopeless. Even if the distribution of the multinomial statement can be derived for some class of f and g functions there is no guarantee that these choices of functions will produce a useful bound on the reliability. The right hand side of the confidence interval statement is multiplicative and hence both

$P_m[g \geq f|A]$ and $P[Z \geq g|A]$ must be capable of simultaneously yielding high numerical values at the sample point. This of course places further restrictions on the class of functions available.

2. BAYESIAN VIEWPOINT

In both previously cited works [1, 3] the probabilities of failure for each of the various modes are considered unknown parameters. In certain cases it may be feasible to assume a prior distribution on the failure modes themselves. As a specific example if the system consists of k components in logical series one failure mode may be associated with each component. Since earlier developmental history should be available on each component it is reasonable to assume that something is known about the reliability of each. On the other hand if each mode of failure is to be some type of failure (electrical, mechanical, etc.), it still may be reasonable to assume that there exists some prior knowledge of these failure rates.

In considering priors in general there are several requirements that must be met. The marginal range of each of the random variables must be the interval zero to one. Each of the component reliabilities will be close to one; hence the distribution chosen for a failure mode must be capable of lumping a good percentage of its probability near zero. The prior should lead to easy computation if possible or at least be tractable.

Three forms of prior distributions were considered.

$$(1) \quad f_{Q_1, Q_2, \dots, Q_k}(q_1, q_2, \dots, q_k) = \begin{cases} G_1 & 0 \leq q_i \leq b \leq 1/k \quad i=1, 2, \dots, k \\ 0 & \text{otherwise} \end{cases}$$

$$(2) \quad f_{Q_1, Q_2, \dots, Q_k}(q_1, q_2, \dots, q_k) = \begin{cases} G_2 q_1^m q_2^m \dots q_k^m & 0 \leq q_i \leq b \leq 1/k \quad i=1, \dots, k \\ 0 & \text{otherwise} \end{cases}$$

$$(3) \quad f_{Q_0, Q_1, \dots, Q_k}(q_0, q_1, \dots, q_k) = G_3 q_0^{m_0-1} q_1^{m_1-1} \dots q_k^{m_k-1}$$

$$\text{where } \sum_{i=0}^k q_i = 1$$

In all three cases q_0 represents the probability of a successful mission and q_i represents the probability of the i^{th} failure mode as before. The G_i 's are **appropriate** constants.

The first two priors assume independence between the various failure modes. Unfortunately the posterior for each of these distributions is difficult to demonstrate. The k fold integral appearing in the denominator of each posterior is non-integrable and numerical techniques would have to be resorted to if it were to be evaluated. This was not deemed worthwhile since the third form of the prior is actually a more general representation than the first two. It is true that (3) disallows independence between failure modes but no strong argument can be made for this independence. The last prior will be discussed more thoroughly and developed in the next section.

3. PRIOR DISTRIBUTION OF COMPONENT FAILURES

If we consider a series of N tests as before with q_0 as the probability of a success on any given trial and q_i , $i=1,2,\dots,k$, as the probability of a failure by the i^{th} failure mode, then the probability of observing N_0 successes and N_i failures, $i=1,2,\dots,k$, in the respective failure modes will follow the multinomial probability law.

Thus,

$$P[N_0, N_1, \dots, N_k | q_0, q_1, \dots, q_k] = \frac{N!}{\prod_{i=0}^k N_i!} q_0^{N_0} q_1^{N_1} \dots q_k^{N_k}$$

$$N_i = 0, 1, 2, \dots, N \quad \sum_{i=0}^k N_i = N$$

Raiffa and Schlaifer¹ suggest using a prior of the same functional form as the distribution of the sample. This procedure leads to prior (3) depicted in the last section. The advantages of this choice for a prior will become apparent in the succeeding development.

Thus,

$$P[q_0, q_1, \dots, q_k] = G q_0^{m_0-1} q_1^{m_1-1} \dots q_k^{m_k-1}$$

$$\text{where } \sum_{i=0}^k q_i = 1 \quad q_i \geq 0$$

shall be the chosen prior. Making use of the usual fact that a density integrated over its range must equal one allows us to evaluate G .

¹Raiffa, H. and Schlaifer, R., Applied Statistical Decision Theory (Boston: Graduate School of Business Administration Harvard University, 1961) p. 47-49.

Hence

$$\begin{aligned} \frac{1}{G} &= \int_{q_1} \dots \int_{q_k} (1 - \sum_{i=1}^k q_i)^{m_0-1} q_1^{m_1-1} \dots q_k^{m_k-1} dq_1 \dots dq_k \\ &= \int_{q_1} \dots \int_{q_{k-1}} q_1^{m_1-1} \dots q_{k-1}^{m_{k-1}-1} dq_1 \dots dq_{k-1} \\ &\quad \int_{q_{k=0}}^{1 - \sum_{i=1}^{k-1} q_i} (1 - \sum_{i=1}^k q_i)^{m_0-1} q_k^{m_k-1} dq_k \end{aligned}$$

Now letting $x_k = \frac{q_k}{1 - \sum_{i=1}^{k-1} q_i}$

$$\frac{1}{G} = \int_{q_1} \dots \int_{q_{k-1}} (1 - \sum_{i=1}^{k-1} q_i)^{m_k+m_0-1} q_1^{m_1-1} \dots q_{k-1}^{m_{k-1}-1} dq_1 \dots dq_{k-1}$$

$$= \int_{q_{k=0}}^1 x_k^{m_k-1} (1 - x_k)^{m_0-1} dx_k$$

It may be noted that the last integral in the above expression is just a Beta integral with parameters m_k and m_0 integrated over its range.

Hence:

$$\int_0^1 x_k^{m_k-1} (1 - x_k)^{m_0-1} dx_k = B(m_k, m_0)$$

By a continued iteration of the above procedure it may be shown that;

$$\frac{1}{G} = B(m_k, m_o) B(m_{k-1}, m_o + m_k) B(m_{k-2}, m_o + m_k + m_{k-1}) \dots$$

$$\dots B(m_1, m_o + m_k + m_{k-1} + \dots + m_2)$$

Noting that

$$B(m_k, m_o) = \frac{\Gamma(m_k) \Gamma(m_o)}{\Gamma(m_k + m_o)}$$

$$G = \frac{\Gamma(m_o + m_1 + \dots + m_k)}{\Gamma(m_o) \Gamma(m_1) \dots \Gamma(m_k)} = \frac{\Gamma\left(\sum_{i=0}^k m_i\right)}{\prod_{i=0}^k \Gamma(m_i)}$$

Hence the complete prior is given by:

$$f_{Q_0, Q_1, \dots, Q_k}(q_0, q_1, \dots, q_k) = \frac{\Gamma\left(\sum_{i=0}^k m_i\right)}{\prod_{i=0}^k \Gamma(m_i)} q_0^{m_o-1} q_1^{m_1-1} \dots q_k^{m_k-1}$$

$$\text{where } \sum_{i=0}^k q_i = 1 \text{ and } q_i \geq 0 \quad i = 0, 1, \dots, k$$

Several properties of this prior will be needed at a later point; for convenience they will be derived or substantiated at this time.

(a) Marginal Distribution of the q_i 's

Consider

$$f_{Q_1}(q_1) = \frac{\Gamma\left(\sum_{i=0}^k m_i\right)}{\prod_{i=0}^k \Gamma(m_i)} \int_{q_2} \cdots \int_{q_k} q_0^{m_0-1} q_1^{m_1-1} \cdots q_k^{m_k-1} dq_2 \cdots dq_k$$

Iterating this integral as in the previous case;

$$f_{Q_1}(q_1) = \frac{\Gamma\left(\sum_{i=0}^k m_i\right)}{\prod_{i=0}^k \Gamma(m_i)} B(m_k, m_0) \cdots B(m_2, m_0 + m_k + \cdots + m_4 + m_3) \\ \times q_1^{m_1-1} (1-q_1)^{m_0+m_k+\cdots+m_3+m_2-1}$$

Letting $\sum_{i=0}^k m_i = m$

$$f_{Q_1}(q_1) = \frac{\Gamma(m)}{\Gamma(m_1) \Gamma(m-m_1)} q_1^{m_1-1} (1-q_1)^{m-m_1-1} \\ = \frac{1}{B(m_1, m-m_1)} q_1^{m_1-1} (1-q_1)^{m-m_1-1} \quad 0 \leq q_1 \leq 1 \\ = 0 \quad \text{otherwise}$$

By symmetry arguments the i^{th} marginal is of the same form, therefore;

$$f_{Q_i}(q) = \frac{1}{B(m_i, m - m_i)} q^{m_i - 1} (1 - q)^{m - m_i - 1}$$

$$= f_B(q, m_i, m - m_i) \quad 0 \leq q \leq 1$$

Thus the marginals of the prior are Beta distributed with the appropriate parameters as shown. For this reason and for convenience in notation the prior itself will be referred to as a multivariate Beta.

(b) Expected Value of the Q_i 's with Respect to the Prior Marginals

$$\begin{aligned} E(Q_i) &= \int_0^1 q f_{Q_i}(q) dq \\ &= \frac{1}{B(m_i, m - m_i)} \int_0^1 q^{m_i} (1 - q)^{m - m_i - 1} dq \end{aligned}$$

Hence

$$E(Q_i) = \frac{B(m_i + 1, m - m_i)}{B(m_i, m - m_i)} = \frac{m_i}{m}$$

(c) Variance of the Q_i 's

$$\begin{aligned} E(Q_i^2) &= \frac{1}{B(m_i, m - m_i)} \int_0^1 q^{m_i + 1} (1 - q)^{m - m_i - 1} dq \\ &= \frac{m_i (m_i + 1)}{m(1 + m)} \end{aligned}$$

Thus

$$\begin{aligned}\text{Var}(Q_i) &= E(Q_i^2) - E^2(Q_i) \\ &= \frac{m_i(m - m_i)}{m^2(1 + m)}\end{aligned}$$

(d) Joint Density of Q_i, Q_j

$$f_{Q_i Q_j}(q_i, q_j) = \frac{\Gamma(m)}{\Gamma(m_i) \Gamma(m_j) \Gamma(m - m_i - m_j)} q_i^{m_i - 1} q_j^{m_j - 1} (1 - q_i - q_j)^{m - m_i - m_j - 1}$$

$$q_i + q_j \leq 1 \quad q_i \geq 0 \quad q_j \geq 0$$

(e) Covariance of Q_i, Q_j

$$\begin{aligned}E(Q_i Q_j) &= \frac{\Gamma(m)}{\Gamma(m_i) \Gamma(m_j) \Gamma(m - m_i - m_j)} \int_{q_i=0}^1 \int_{q_j=0}^{1-q_i} q_i^{m_i} q_j^{m_j} (1 - q_i - q_j)^{m - m_i - m_j - 1} dq_i dq_j \\ &= \frac{\Gamma(m)}{\Gamma(m_i) \Gamma(m_j) \Gamma(m - m_i - m_j)} \int_{q_i=0}^1 \int_{q_j=0}^{1-q_i} q_i^{m_i} q_j^{m_j} (1 - q_i - q_j)^{m - m_i - m_j - 1} dq_i dq_j\end{aligned}$$

Let

$$x = \frac{q_j}{1 - q_i}$$

Then integrating as before

$$E(Q_i Q_j) = \frac{m_i m_j}{m(1 + m)}$$

And

$$\begin{aligned}\text{Cov } (Q_i Q_j) &= E(Q_i Q_j) - E(Q_i) E(Q_j) \\&= \frac{m_i m_j}{m(1+m)} - \frac{m_i m_j}{m^2} \\&= - \frac{m_i m_j}{m^2(1+m)}\end{aligned}$$

4. POSTERIOR DISTRIBUTION OF THE PARAMETERS

In the process described previously the distribution of the sample was given by

$$P[N_0, N_1, \dots, N_k | q_0 \dots q_k] = \frac{N!}{N_0! \dots N_k!} q_0^{N_0} q_1^{N_1} \dots q_k^{N_k}$$

Utilizing the prior described in section three and applying Bayes rule we get;

$$P[q_0 \dots q_k | N_0, \dots, N_k] = \frac{q_0^{N_0+m_0-1} q_1^{N_1+m_1-1} \dots q_k^{N_k+m_k-1}}{\int_{q_1} \dots \int_{q_k} q_1^{N_1+m_1-1} \dots q_k^{N_k+m_k-1} [1 - \sum_{i=1}^k q_i]^{N_0+m_0-1} dq_1 \dots dq_k}$$

Noting the similarity between this integral and that previously seen in section three gives

$$P[q_0 \dots q_k | N_0, \dots, N_k] = \frac{\Gamma(N+m)}{\Gamma(N_0+m_0) \dots \Gamma(N_k+m_k)} q_0^{N_0+m_0-1} \dots q_k^{N_k+m_k-1}$$

$$\text{where } N = \sum_{i=0}^k N_i \quad \text{and } m = \sum_{i=0}^k m_i \quad \text{as before.}$$

Except for suitable changes in parameters this distribution can be recognized as the multivariate Beta of the previous section. This facility in handling the posterior is, of course, one of the reasons for choosing this particular distribution for a prior. Of course a prior must meet more requirements than just facility in use. The

multivariate Beta in question does satisfy the intuitive demands set forth earlier. A logical method for selecting the parameters of the prior (m_i 's) will be given in Appendix I. The assumption of this prior is equivalent to a form of "linear squashing" described by Good¹ in his work on the estimation of probabilities.

As has been noted above the posterior distribution is itself a Beta, hence the properties of the posterior may be stated by analogy from the previous results. For future convenience these properties will be listed below.

(a) Posterior Marginal

$$f_{Q_i}(q | N_o, \dots, N_k) = \frac{1}{B(N_i + m_i, N + m - (N_i + m_i))} q^{N_i + m_i - 1} (1 - q)^{N + m - (N_i + m_i) - 1}$$

(b) Marginal Mean

$$E(Q_i | N_o, \dots, N_k) = \frac{N_i + m_i}{N + m}$$

(c) Marginal Variance

$$\text{Var}(Q_i | N_o, \dots, N_k) = \frac{(N + m) (N_i + m_i) - (N_i + m_i)^2}{(N + m)^2 (1 + N + m)}$$

(d) Conditional Covariance of Q_i, Q_j

$$\text{Cov}(Q_i Q_j | N_o, \dots, N_k) = - \frac{(N_i + m_i) (N_j + m_j)}{(N + m)^2 (1 + N + m)}$$

¹ Good, I. J., The Estimation of Probabilities, an Essay on Modern Bayesian Methods (Cambridge: Research Monograph No. 30, The M.I.T. Press, 1965) p. 24.

5. RELIABILITY

Corcoran, Weingarten and Zehna [1] define the current measure of reliability, p^* , that was stated in section one of this paper. The estimation problem is considered to have taken place prior to the observation of the N tests. Averaging over all possible outcomes of the N tests to produce the "mean reliability" is justified on this basis. Larson [3] regards this of interest in the early phases of development before testing can take place but points out that the final reliability is actually a function of the outcomes of the N tests. This is true because whether or not a corrective attempt is made depends on observing the given failure mode. For this reason he develops the conditional distribution of true reliability (conditioned on the outcomes of the N tests) and shows that p^* is in fact the mean of this conditional distribution. This poses the question of how to properly interpret $E(p^*)$. Since p^* is the mean of the conditional reliability it would seem unreasonable to average over all possible outcomes of the N tests (some of which are known not to have occurred) to obtain $E(p^*)$. Larson considers it more reasonable to attempt to estimate p^* rather than $E(p^*)$.

There would seem to be other avenues of approach open. Consider the original question of attempting to make probability statements about a complex system after some period of testing and applying corrective action. The discreteness of the conditional reliability function could be interpreted as an anomaly arising from simplification in the logical statement of the problem. The uncertainty in the actual reliability arises from two sources: the uncertainty in the failure modes themselves and the uncertainty of correcting a type of failure if it occurs. In

some cases it may be feasible to treat the uncertainty in the failure modes by assuming a prior distribution on the q_i 's. Partial justification for this procedure was given in section two.

It does appear attractive to treat the reliability as conditioned on the outcomes of the N tests. At least at that point in time we know which modes of failure it is possible to remove. With the above considerations in mind the following scheme is proposed. Let N tests be conducted, the results of which are N_0 successes and N_i failures in the respective failure modes, $i = 1, 2, \dots, k$. The probability of a success is q_0 , the probability of the failure of the i^{th} mode is q_i , where

$$\sum_{i=0}^k q_i = 1 \qquad \sum_{i=0}^k N_i = N$$

The maximum number of failure modes that can occur is k ; it is not necessary that each mode be observed however, and a sample will in general demonstrate less than k failure types. If we let a_i be the probability of correcting the i^{th} failure mode given that it is observed and b_i be the unconditional probability of correcting each failure mode it follows that

$$b_i = \begin{cases} 0 & N_i = 0 \\ a_i & N_i > 0 \end{cases}$$

The following is stated as an intuitive measure of the reliability after corrective action. This is of course the same measure defined by Corcoran, Weingarten and Zehna [1]. The interpretation given to it here, however, differs from that of the above paper. The actual sampling is considered to have taken place and R is conditioned on this

outcome through the b_i 's.

$$\begin{aligned}
 R &= q_0 + \sum_{i=1}^k b_i q_i \\
 &= 1 - \sum_{i=1}^k q_i + \sum_{i=1}^k b_i q_i \\
 &= 1 + \sum_{i=1}^k (b_i - 1) q_i
 \end{aligned}$$

If we are willing to assume a distribution on the q_i 's then R is a function of random variables and is hence itself a random variable. The remainder of this study will be devoted to the development of the properties of this measure of reliability with respect to the posterior distribution developed in section four. R is the weighted sum of dependent Beta distributed random variables. It is not possible to obtain the convolution of independent Beta's, hence there is little or no hope of deriving the distribution of R .

Now

$$E(R) = 1 + \sum_{i=1}^k (b_i - 1) E(q_i)$$

Noting from section four

$$\begin{aligned}
 E(q_i) &= \frac{N_i + m_i}{N + m} \\
 E(R) &= 1 + \sum_{i=1}^k (b_i - 1) \frac{N_i + m_i}{N + m}
 \end{aligned}$$

Using the variance and covariance formulation developed in section four,
let

$$\sigma_{ij} = - \frac{(N_i + m_i)(N_j + m_j)}{(N + m)^2 (N + m + 1)}$$

$$\sigma_{ii} = \frac{(N + m)(N_i + m_i) - (N_i + m_i)^2}{(N + m)^2 (N + m + 1)}$$

Now let V equal the matrix

$$\begin{array}{l} || \sigma_{ij} || \quad i = 1, 2, \dots, k \\ \quad \quad \quad j = 1, 2, \dots, k \end{array}$$

and $\underline{a}'$ equal the vector

$$(b_1 - 1, b_2 - 1, \dots, b_k - 1)$$

Then the variance of R is the quadratic form

$$\text{Var}(R) = \underline{a}' V \underline{a}$$

Thus although the distribution of R cannot be shown its mean and variance are known. These parameters are functions of the N_i 's and hence are uniquely determined for each sample point. In an informal sense the mean and variance of a random variable can be said to "estimate" the random variable and hence $E(R)$ provides some indication of the corrected value of the reliability. In the above case it may be noted that the variance of R tends to zero as N tends to infinity; hence the density of R "lumps" at the mean as N increases. In this sense $E(R)$ may be said to asymptotically estimate R as N increases. This fact may be of little use however since the N 's of interest can be expected to be small.

It is feasible that the higher moments of the distribution could be obtained. While these moments may be complex they should be functions of the N 's and hence capable of being evaluated. Obtaining all of the moments would of course be equivalent to obtaining the actual distribution.

One last avenue of approach may be worthy of mention. It was initially hoped that the formulation in this study would lead to the ability to make confidence interval statements about the reliability. Except for the loose bound obtainable through Chebychev's inequality this has not been realized. The question of the existence of some limit theorem has not been thoroughly investigated. This does not appear promising but it may warrant some further consideration.

BIBLIOGRAPHY

1. Corcoran, W. J., Weingarten, H. and Zehna, P. W. Estimating Reliability after Corrective Action. Management Science, vol. 10, No. 4, July 1964: pp. 786-795.
2. Good, I. J. The Estimation of Probabilities, An Essay on Modern Bayesian Methods. The M.I.T. Press, 1965.
3. Larson, H. J. Conditional Distribution of True Reliability after Corrective Action. U. S. Naval Postgraduate School. Technical Report/Research Paper No. 61, January 1966.
4. Raiffa, H. and Schlaifer, R., Applied Statistical Decision Theory. Graduate School of Business Administration Harvard University, 1961.
5. Silver, E. A., Markovian Decision Processes with Uncertain Transition Probabilities or Rewards. Operations Research Center, Massachusetts Institute of Technology. Interim Technical Report No. 1, August 1963.

APPENDIX I

SELECTION OF PRIOR PARAMETERS

One possible method of selecting the parameters of a prior distribution would be to fit a set of prior means and variances to the distribution. With the multivariate Beta in question however there are insufficient parameters $(k + 1)$ to do this. The following method suggested by Silver¹ is given.

We use the last k parameters to fit the means, then the remaining parameter, m_0 , is used to obtain a least squares fit to the prior variances.

Let

$$(1) \quad m = \sum_{i=0}^k m_i$$

Then

$$E(q_i) = \frac{m_i}{m} \quad i = 1, 2, \dots, k$$

$$(2) \quad \text{Var}(q_i) = \frac{m_i(m - m_i)}{m(1 + m)} \quad i = 1, 2, \dots, k$$

Substituting equation (1) into equation (2)

$$\text{Var}(q_i) = \frac{E(q_i) [1 - E(q_i)]}{m + 1}$$

Let $\overline{\text{Var}(q_i)}$ be the prior estimated variance of q_i .

¹ Silver, E. A., Markovian Decisions Processes with Uncertain Transition Probabilities or Rewards (M.I.T., Interim Technical Report No. 1, 1963) Appendix C, pp. 177-179.

Hence we want to choose m to minimize

$$SS = \sum_{i=1}^k (\overline{\text{Var}(q_i)} - \text{Var}(q_i))^2$$

$$SS = \sum_{i=1}^k \overline{\text{Var}(q_i)} - \left[\frac{E(q_i) [1 - E(q_i)]}{m + 1} \right]^2$$

Setting $d(SS)/dm$ equal to zero we get the following for a least squares value of m equal to $\hat{m}$.

$$\hat{m} = \frac{\sum_{i=1}^k [E(q_i)]^2 [1 - E(q_i)]^2}{\sum_{i=1}^k \overline{\text{Var}(q_i)} E(q_i) [1 - E(q_i)]} - 1$$

and hence

$$m_i = \hat{m} E(q_i) \quad i = 1, 2, \dots, k$$

INITIAL DISTRIBUTION LIST

	No. Copies
1. Defense Documentation Center Cameron Station Alexandria, Virginia 22314	20
2. Library U. S. Naval Postgraduate School Monterey, California	2
3. Professor H. J. Larson Department of Operations Analysis U. S. Naval Postgraduate School Monterey, California	1
4. LCDR Charles M. Earnest VA-125 Naval Air Station Lemoore, California	1

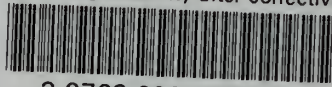
DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) UNITED STATES NAVAL POSTGRADUATE SCHOOL MONTEREY, CALIFORNIA		2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	
		2b. GROUP ----	
3. REPORT TITLE ESTIMATING RELIABILITY AFTER CORRECTIVE ACTION: A BAYESIAN VIEWPOINT			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Master's Thesis			
5. AUTHOR(S) (Last name, first name, initial) Earnest, Charles M.			
6. REPORT DATE May 1966		7a. TOTAL NO. OF PAGES 30	7b. NO. OF REFS 5
8a. CONTRACT OR GRANT NO.		9a. ORIGINATOR'S REPORT NUMBER(S)	
b. PROJECT NO.			
c.		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.			
10. AVAILABILITY/LIMITATION NOTICES Qualified requesters may obtain copies of this report from DDC.			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY U. S. Naval Postgraduate School Monterey, California	
13. ABSTRACT A complex system is considered in its latter stages of development. N mission trials have been observed, each resulting in a success or a failure. Each failure occurs in one of k failure modes. For each failure mode that is observed action is taken to attempt to correct that type of failure. The probabilities of correcting the various failure modes are known. After corrective action is completed attempts to estimate the current reliability, without further sampling, are made. A brief historical summary of this problem to date is given. Justification for assuming a prior distribution on the failure modes is discussed and the posterior distribution of the parameters is developed. An intuitive measure of the current reliability is stated and certain properties of this random variable are developed.			

thesE118

Estimating reliability after corrective



3 2768 000 98742 4

DUDLEY KNOX LIBRARY